Performance Assessment of Universal Machine Learning Interatomic Potentials: Challenges and Directions for Materials' Surfaces

Bruno Focassio[1] , Luis Paulo M Freitas[1] , **Gabriel R. Schleder [1,2]**

[1] Brazilian Nanotechnology National Laboratory (LNNano/CNPEM), 13083-100, Campinas, SP, Brazil
[2] John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA

gabriel.schleder@lnnano.cnpem.br

Machine learning interatomic potentials (MLIPs) are one of the main techniques in the materials science toolbox, able to bridge ab initio accuracy with the computational efficiency of classical force fields. This allows simulations ranging from atoms, molecules, and biosystems, to solid and bulk materials, surfaces, nanomaterials, and their interfaces and complex interactions.

A recent class of advanced MLIPs, which use equivariant representations and deep graph neural networks, is known as universal models [1-4]. These models are proposed as foundational models suitable for any system, covering most elements from the periodic table. Current universal MLIPs (UIPs) have been trained with the largest consistent dataset available nowadays. However, these are composed mostly of bulk materials' DFT calculations.
In this presentation, we assess the universality of all openly available UIPs, namely MACE, CHGNet, and M3GNet, in a representative task of generalization: calculation of surface energies. We find that the out-of-the-box foundational models have significant shortcomings in this task (Figs. 1-2), with errors correlated to the total energy of surface simulations, having an out-of-domain distance from the training dataset. Our results [5] show that while UIPs are an efficient starting point for fine-tuning specialized models (Fig. 3), we envision the potential of increasing the coverage of the materials space towards universal training datasets for MLIPs.

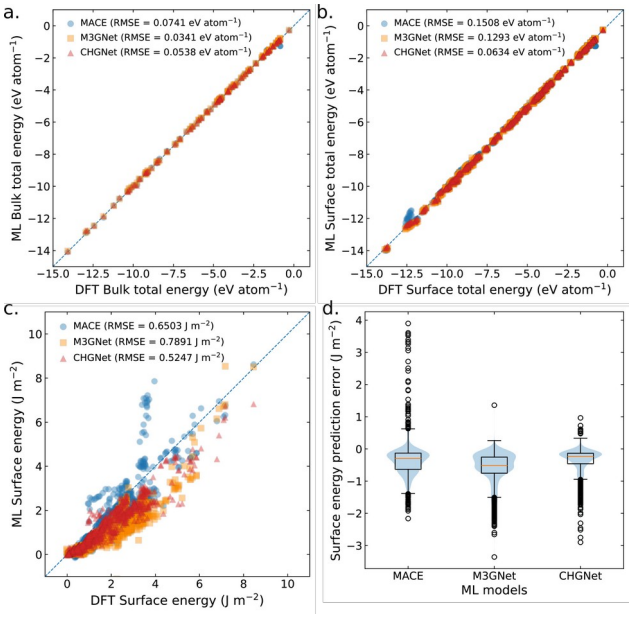

Figure 1. Performance assessment of the universal interatomic potentials over the surfaces dataset. (a) Parity plot for the total energy per atom of the bulk systems that gave origin to the surfaces of the dataset. (b) Parity plot for the total energy per atom of the surfaces within the dataset. (c) Parity plot for the surface energy (γ σ hkl) for the surfaces within the dataset. For (a), (b), and (c), the dashed line marks the x = y line. (d) Boxplot and violin plot for the error in the prediction of the surface energy (γ σ hkl) from the three universal interatomic potentials evaluated. The horizontal lines in the middle of the boxes mark the medians. The boxes are plotted from the first to the third quartile. The whiskers extend to 1.5 times the interquartile range from the third (above) and first (below) quartile and empty circles mark the outliers (data points outside the whiskers).

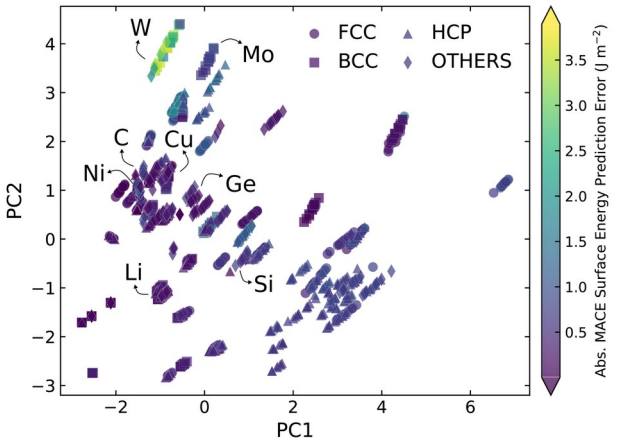

Figure 2. Kernel PCA (kPCA) map of the surfaces dataset, where each point corresponds to a structure and distances represent differences in input features. kPCA map with MACE representation and REMatch kernel, colored by the absolute surface energy error given by the MACE model, with selected elements highlighted. Markers represent the lattice structure of the bulk that generated each surface, with FCC, BCC, and HCP crystalline systems highlighted.

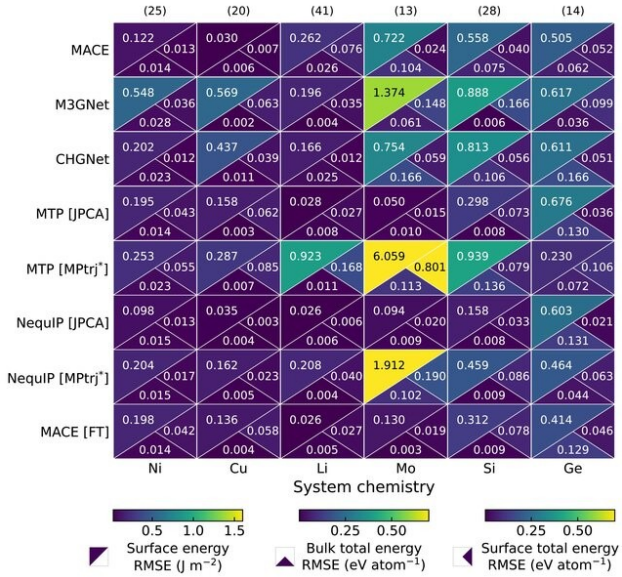

Figure 3. Performance comparison between universal, specialized, and universal fine-tuned interatomic potentials. The root mean squared errors (RMSE) of the universal interatomic potentials are compared in relation to the surface chemistry, for selected elements. The upper triangle is the surface energy (γ σ hkl) RMSE, the lower leftmost triangle is the bulk total energy per atom RMSE and the lower rightmost triangle is the surface total energy per atom RMSE. The numbers in parenthesis are the number of surface structures evaluated for each chemical element.

References

[1] C. Chen and S. P. Ong, A universal graph deep learning interatomic potential for the periodic table, Nature Computational Science 2, 718–728 (2022).

[2] B. Deng, P. Zhong, K. Jun, J. Riebesell, K. Han, C. J. Bartel, and G. Ceder, Chgnet as a pretrained universal neural network potential for charge-informed atomistic modelling, Nature Machine Intelligence 5, 1031–1041 (2023).

[3] I. Batatia, P. Benner, Y. Chiang, A. M. Elena, D. P. Kovács, J. Riebesell, X. R. Advincula, M. Asta, W. J. Baldwin, N. Bernstein, A. Bhowmik, S. M. Blau, V. Cărare, J. P. Darby, S. De, F. Della Pia, V. L. Deringer, R. Elijošius, Z. El-Machachi, E. Fako, A. C. Ferrari, A. Genreith-Schriever, J. George, R. E. A. Goodall, C. P. Grey, S. Han, W. Handley, H. H. Heenen, K. Hermansson, C. Holm, J. Jaafar, S. Hofmann, K. S. Jakob, H. Jung, V. Kapil, A. D. Kaplan, N. Karimitari, N. Kroupa, J. Kullgren, M. C. Kuner, D. Kuryla, G. Liepuoniute, J. T. Margraf, I.-B. Magdău, A. Michaelides, J. H. Moore, A. A. Naik, S. P. Niblett, S. W. Norwood, N. O'Neill, C. Ortner, K. A. Persson, K. Reuter, A. S. Rosen, L. L. Schaaf, C. Schran, E. Sivonxay, T. K. Stenczel, V. Svahn, C. Sutton, C. van der Oord, E. Varga-Umbrich, T. Vegge, M. Vondrák, Y. Wang, W. C. Witt, F. Zills, and G. Csányi, A foundation model for atomistic materials chemistry (2024).

[4] A. Merchant, S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon, and E. D. Cubuk, Scaling deep learning for materials discovery, Nature 624, 80–85 (2023).

[5] Focassio, Bruno ; Mezzina Freitas, Luis Paulo ; Schleder, Gabriel R.. Performance Assessment of Universal Machine Learning Interatomic Potentials: Challenges and Directions for Materials' Surfaces. arXiv eprint (2024), DOI: 10.48550/arXiv.2403.04217